



UNIVERSITY OF  
ILLINOIS LIBRARY  
AT URBANA-CHAMPAIGN  
STACKS

B24E 2-3

**CENTRAL CIRCULATION BOOKSTACKS**

The person charging this material is responsible for its renewal or its return to the library from which it was borrowed on or before the **Latest Date** stamped below. **You may be charged a minimum fee of \$75.00 for each lost book.**

**Theft, mutilation, and underlining of books are reasons for disciplinary action and may result in dismissal from the University.**

**TO RENEW CALL TELEPHONE CENTER, 333-8400**

**UNIVERSITY OF ILLINOIS LIBRARY AT URBANA-CHAMPAIGN**

JUL 25 1995

When renewing by phone, write new due date below previous due date.

L162

Digitized by the Internet Archive  
in 2011 with funding from  
University of Illinois Urbana-Champaign

<http://www.archive.org/details/convexoptimizati407blai>

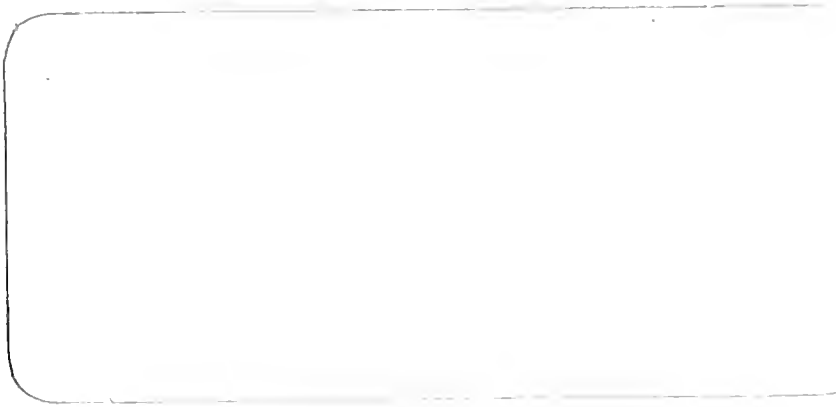
## Faculty Working Papers

CONVEX OPTIMIZATION AND LAGRANGE MULTIPLIERS

Charles E. Blair

#407

**College of Commerce and Business Administration**  
**University of Illinois at Urbana-Champaign**



FACULTY WORKING PAPERS

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

June 8, 1977

CONVEX OPTIMIZATION AND LAGRANGE MULTIPLIERS

Charles E. Blair

#407

1.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 2.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 3.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 4.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 5.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 6.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 7.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 8.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 9.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 10.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

### Section 1.1

1.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 2.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 3.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 4.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 5.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 6.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 7.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 8.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 9.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$   
 10.  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$



Convex Optimization and Lagrange Multipliers

by

Charles E. Blair

Dept. of Business Administration

June 8, 1977

This work was supported by a grant from Investors in Business Education,  
University of Illinois.

Abstract

We show how the duality theorem of linear programming can be used to  
prove several results on general convex optimization.



Let  $f, g_1, \dots, g_k$  be convex functions defined on a convex subset  $S$  of a vector space. Let  $T = \{x \in S \mid g_i(x) \leq 0 \ 1 \leq i \leq k\}$ . We assume throughout that  $T$  is non-empty.

We use linear programming theory to explore the relationship between the problem of minimizing  $f(x)$   $x \in T$  and the Lagrange dual problem of minimizing  $f(x) + \sum \lambda_i g_i(x)$   $x \in S$  for suitable  $\lambda_i \geq 0$ . This is motivated by the work of Duffin [1, 2, 3]. The main tool we shall need is a version of the duality theorem of linear programming [5, theorems 1.1.9 and 1.7.13].

Lemma: Let  $A = \{x \mid Bx \geq b \text{ and } Cx = c\}$ . If  $A$  is empty, there are  $U \geq 0$  and  $V$  such that  $U^t B + V^t C = \vec{0}$  and  $Ub + Vc > 0$ . If  $A$  is non-empty and every  $x \in A$  satisfies  $dx \geq e$ , then there are  $U \geq 0$  and  $V$  such that  $U^t B + V^t C = d$  and  $Ub + Vc \geq e$ .

We will not require separating hyperplane theorems or results from semi-infinite programming.

Theorem 1:  $f(x) \geq L$  for every  $x \in T$  if and only if for every finite FCS there are  $\lambda_1, \dots, \lambda_k$  such that  $f(x) + \sum \lambda_i g_i(x) \geq L$  for every  $x \in F$ .

Proof: The "if" part is immediate. For the "only if" part it suffices to prove the result for those finite  $F$  which contain members of  $T$ . Let  $F = \{y_1, \dots, y_N\}$   $y_i \in T$ . For such  $F$  the system of equations and inequalities in unknowns  $\theta_1, \dots, \theta_N$

$$(D) \quad \begin{aligned} & \sum_{i=1}^N \theta_i = 1 \\ & \sum_{i=1}^N \theta_i g_j(y_i) \leq 0 \quad 1 \leq j \leq k \\ & \theta_i \geq 0 \end{aligned}$$

has the solution  $\theta_1 = 1$ . By convexity, if  $\theta_i$  is a solution to (D),  $\sum \theta_i y_i \in T$ .



Since we are assuming  $f(x) \geq L$  for  $x \in T$ , every solution to (D) must satisfy  $\sum_{i=1}^k \lambda_i f(y_i) \geq L$ . By the lemma there are  $\lambda_i \geq 0$  and  $\gamma$  such that

$$\gamma + \sum_{j=1}^k \lambda_j (-g_j(y_i)) \leq f(y_i) \quad 1 \leq i \leq N$$

$$\gamma \geq L$$

so  $f(y_i) + \sum_{j=1}^k \lambda_j g_j(y_i) \geq L$  for every  $y_i \in F$ . Q.E.D.

Theorem 1 may be used to prove many of the standard results on convex optimization. As an example we prove the Kuhn-Tucker theorem.

Corollary: Suppose  $f(x) \geq L$  for  $x \in T$  and that there is a  $y$  for which  $g_i(y) < 0$   $1 \leq i \leq k$ . Then there are  $\lambda_1, \dots, \lambda_k$  such that  $f(x) + \sum_{i=1}^k \lambda_i g_i(x) \geq L$  for  $x \in S$ .

Proof: Let  $\delta = \max \{g_i(y)\}$ . For  $x \in S$  let  $A_x = \{(\lambda_1, \dots, \lambda_k) \mid \lambda_i \geq 0, f(x) + \sum_{i=1}^k \lambda_i g_i(x) \geq L, \text{ and } -\delta(\sum_{i=1}^k \lambda_i) \leq f(y) - L\}$ . For each  $x \in S$ ,  $A_x$  is compact. If HCS is finite we may use theorem 1 with  $F = H \cup \{y\}$  to show  $\bigcap_{x \in H} A_x$  is non-empty. Therefore,  $\bigcap_{x \in S} A_x$  is non-empty, so suitable Lagrange multipliers exist. Q.E.D.

Arguments of this kind can also be used to give information about when "duality gaps" occur.

Theorem 2. There are no  $\lambda_i$  such that  $f(x) + \sum_{i=1}^k \lambda_i g_i(x) \geq L$  for every  $x \in S$  if and only if, for every  $N > 0$ , there is an  $x \in S$  such that  $f(x) < L - N(\max_{1 \leq i \leq k} g_i(x))$ .

Proof: Take any  $\lambda_i \geq 0$  and suppose  $x$  exists with the desired property for  $N = \sum \lambda_i$ . Then  $f(x) + \sum_{i=1}^k \lambda_i g_i(x) \leq f(x) + (\max_{i=1}^k g_i(x)) (\sum \lambda_i) < L$ . So no suitable  $\lambda_i$  exist.

Conversely, suppose there are no suitable  $\lambda_i$ . For any  $N > 0$  and any  $F = \{y_1, \dots, y_M\} \subset S$  consider the linear system in unknowns  $\lambda_1, \dots, \lambda_k$ :



$$f(y_i) + \sum_1^k \lambda_j g_j(y_i) \geq L \quad y_i \in F$$

$$(F;N) \quad \sum \lambda_i \leq N$$

$$\lambda_i \geq 0$$

If, for some  $N$ ,  $(F;N)$  had a solution for every finite FCS, a compactness argument similar to that in the corollary to Theorem 1 would yield suitable multipliers  $\lambda_i$ . Since we are assuming such  $\lambda_i$  do not exist, it must be that for every  $N > 0$  there is an  $F$  such that  $(F;N)$  has no solution. By the lemma, if  $(F;N)$  has no solution, there are  $\theta_1, \dots, \theta_M \geq 0$  and  $\gamma > 0$  such that

$$\sum_{i=1}^M \theta_i g_j(y_i) - \gamma \leq 0, \quad 1 \leq j \leq k$$

and  $\sum \theta_i (L - f(y_i)) + \gamma(-N) > 0$ . By scaling, we may assume  $\sum_{i=1}^M \theta_i = 1$ , so that

$$\sum_{i=1}^M \theta_i g_j(y_i) \leq \gamma < \frac{1}{N} (L - \sum \theta_i f(y_i)).$$

If we take  $x = \sum \theta_i y_i$ ,  $\gamma \geq g_j(x)$ ,  $1 \leq j \leq k$  and  $f(x) < L - N\gamma$  follow by convexity of  $f$  and  $g_j$ . Q.E.D.

Corollary: (Compare [1], cor. 5; [2], thm. 3): Let  $h(\epsilon) = \inf\{f(x) \mid g_j(x) \leq \epsilon, 1 \leq j \leq k\}$ . If there are  $\delta > 0$ ,  $L$  such that  $h(x) > L$  for  $0 < x \leq \delta$ , then there are  $\lambda_1, \dots, \lambda_k$  such that  $f(x) + \sum \lambda_i g_i(x) \geq L \quad x \in S$ .

Proof: If there is an  $x$  for which  $g_j(x) < 0$  the existence of suitable  $\lambda_i$  follows from the Kuhn-Tucker theorem,\* so we assume this is not the case.  $h$  is a convex monotone function which, on our assumptions, is defined only for non-negative arguments. For  $\epsilon \geq \delta$ ,  $h(\epsilon) - L > h(\epsilon) - h(0) \geq \epsilon \left(\frac{1}{\delta}\right) (h(\delta) - h(0))$ . Hence for  $x \in S$ ,  $f(x) - L \geq (\max g_i(x)) (\min 0, \frac{1}{\delta} (h(\delta) - h(0)))$ . (Note that our assumptions

---

\* We use the Kuhn-Tucker theorem for brevity. The result could be proved from Theorem 2 alone.





imply  $\max g_j(x) \geq 0$ .) Since the condition given by Theorem 2 fails for  $N = \max(0, \frac{1}{\delta}(h(0) - h(\delta)))$ , suitable  $\lambda_i$  exist.) Q.E.D.

Finally, we use a variation of these techniques to strengthen a recent result of Duffin and Jeroslow [4].

Theorem 3: Let  $S = \mathbb{R}^n$ . Assume that for  $\lambda_i > 0$ ,  $f(x) + \sum \lambda_i g_i(x) \geq L$ ,  $x \in S$ . Then there are affine functions  $h_i(x) = a_i x + b_i$  ( $a_i \in \mathbb{R}^n$ ,  $b_i \in \mathbb{R}$ ) such that  $h_i(x) \leq g_i(x)$  and  $f(x) + \sum \lambda_i h_i(x) \geq L$ ,  $x \in S$ .

Proof: For  $y \in S$ , let  $T_y = \{(h_1, \dots, h_k) \mid h_i \text{ affine, } h_i(y) \leq g_i(y), \text{ and } f(y) + \sum \lambda_i h_i(y) \geq L\}$ .

We identify each function  $h_i(x) = a_i x + b_i$  with the ordered pair  $(a_i, b_i) \in \mathbb{R}^{n+1}$ . Thus,  $T_y$  is a closed subset of  $\mathbb{R}^{(n+1)k}$ . We first show that, for any finite FCS  $\bigcap_{y \in F} T_y$  is non-empty. A member of  $\bigcap_{y \in F} T_y$  would be a solution to the linear inequality system in unknowns  $a_1, \dots, a_k; b_1, \dots, b_k$ .

$$\begin{aligned} a_i y + b_i &\leq g_i(y) & 1 \leq i \leq k; & y \in F \\ (E) \quad \sum_{i=1}^k \lambda_i (a_i y + b_i) &\geq M - f(y) & y \in F \end{aligned}$$

By the lemma, (E) has no solution only if there are scalars  $W_{iy}, V_y \geq 0$  such that

$$\begin{aligned} (i) \quad \sum_{y \in F} W_{iy} y &= \lambda_i \sum_{y \in F} V_y y, & 1 \leq i \leq k \\ (ii) \quad \sum_{y \in F} W_{iy} &= \lambda_i \sum_{y \in F} V_y, & 1 \leq i \leq k \\ (iii) \quad \sum_{\substack{1 \leq i \leq k \\ y \in F}} W_{iy} g_i(y) &< \sum_{y \in F} V_y (M - f(y)). \end{aligned}$$

If there were  $W, V$  satisfying (i)-(iii) we could set  $V'_y = V_y / \sum_{y \in F} V_y$  and

$W'_{iy} = W_{iy} / \lambda_i \sum_{y \in F} V_y$  so that (i)-(iii) would be satisfied and, by (ii),  $\sum_{y \in F} V'_y = \sum_{y \in F} W'_{iy} = 1$ ,  $1 \leq i \leq k$ . Condition (i) becomes  $\sum_{y \in F} W'_{iy} y = \sum_{y \in F} V'_y y$ . Condition (iii) becomes  $\sum_{y \in F} V'_y f(y) + \sum_{y \in F} \sum_{i=1}^k W'_{iy} g_i(y) < M$ . If  $z = \sum_{y \in F} W'_{iy} y = \sum_{y \in F} V'_y y$  this implies, by convexity,  $f(z) + \sum_{i=1}^k \lambda_i g_i(z) < M$



which contradicts our assumption about the  $\lambda_i$ . Therefore (E) has solutions for every finite F.

To complete the proof we must show  $\bigcap_{y \in S} T_y$  is non-empty. Let  $e_j = j$ th unit vector. We show that if F contains  $\pm e_j$   $1 \leq j \leq n$  and the zero vector, then  $\bigcap_{y \in F} T_y$  is bounded. Since each  $T_y$  is closed, compactness yields the desired result. For  $1 \leq i \leq k$  we must have  $h_i(\vec{0}) = b_i \leq g_i(0)$ ,  $f(\vec{0}) + \sum \lambda_i b_i \geq L$ ,  $h_i(e_j) = (j$ th component of  $a_i) + b_i \leq g_i(e_j)$ , and  $h_i(-e_j) \leq g_i(-e_j)$ . Since all the  $\lambda_i$  are positive this implies bounds on  $a_i$ ,  $b_i$ . Q.E.D.

Our proof of Theorem 3 works for any convex  $S \subset \mathbb{R}^n$  which includes  $\pm e_j$  and  $\vec{0}$ . By suitable translations, this implies the results for any fully dimensional convex  $S \subset \mathbb{R}^n$ . Further modifications yield the result for arbitrary convex  $S \subset \mathbb{R}^n$ .

I would like to thank Richard Duffin and Robert Jeroslow for their encouragement, and for supplying me "sneak previews" of [2] and [4].



## REFERENCES

1. Duffin, R.J. "Convex Analysis Treated by Linear Programming." Mathematical Programming, 4, pp. 125-43.
2. \_\_\_\_\_. "The Lagrange Multiplier Method for Convex Programming." Proceedings of National Academy of Sciences, 72, pp. 1778-1781.
3. \_\_\_\_\_. "Convex Programming Having Some Linear Constraints." Proceedings of National Academy of Sciences, 74, pp. 26-8.
4. Duffin, R.J. and Jeroslow, R.G. Private communication.
5. Stoer, J. and Witzgall, C. Convexity and Optimization in Finite Dimensions I. Springer-Verlag, 1970.

The following table shows the results of the experiment. The first column is the number of trials, the second column is the number of correct responses, and the third column is the percentage of correct responses. The fourth column is the number of errors, and the fifth column is the percentage of errors. The sixth column is the number of omissions, and the seventh column is the percentage of omissions. The eighth column is the number of correct responses per trial, and the ninth column is the percentage of correct responses per trial. The tenth column is the number of errors per trial, and the eleventh column is the percentage of errors per trial. The twelfth column is the number of omissions per trial, and the thirteenth column is the percentage of omissions per trial.

Trial	Correct	% Correct	Errors	% Errors	Omissions	% Omissions	Correct/Trial	% Correct/Trial	Errors/Trial	% Errors/Trial	Omissions/Trial	% Omissions/Trial
1	1	100	0	0	0	0	1	100	0	0	0	0
2	1	100	0	0	0	0	1	100	0	0	0	0
3	1	100	0	0	0	0	1	100	0	0	0	0
4	1	100	0	0	0	0	1	100	0	0	0	0
5	1	100	0	0	0	0	1	100	0	0	0	0
6	1	100	0	0	0	0	1	100	0	0	0	0
7	1	100	0	0	0	0	1	100	0	0	0	0
8	1	100	0	0	0	0	1	100	0	0	0	0
9	1	100	0	0	0	0	1	100	0	0	0	0
10	1	100	0	0	0	0	1	100	0	0	0	0
11	1	100	0	0	0	0	1	100	0	0	0	0
12	1	100	0	0	0	0	1	100	0	0	0	0
13	1	100	0	0	0	0	1	100	0	0	0	0
14	1	100	0	0	0	0	1	100	0	0	0	0
15	1	100	0	0	0	0	1	100	0	0	0	0
16	1	100	0	0	0	0	1	100	0	0	0	0
17	1	100	0	0	0	0	1	100	0	0	0	0
18	1	100	0	0	0	0	1	100	0	0	0	0
19	1	100	0	0	0	0	1	100	0	0	0	0
20	1	100	0	0	0	0	1	100	0	0	0	0
21	1	100	0	0	0	0	1	100	0	0	0	0
22	1	100	0	0	0	0	1	100	0	0	0	0
23	1	100	0	0	0	0	1	100	0	0	0	0
24	1	100	0	0	0	0	1	100	0	0	0	0
25	1	100	0	0	0	0	1	100	0	0	0	0
26	1	100	0	0	0	0	1	100	0	0	0	0
27	1	100	0	0	0	0	1	100	0	0	0	0
28	1	100	0	0	0	0	1	100	0	0	0	0
29	1	100	0	0	0	0	1	100	0	0	0	0
30	1	100	0	0	0	0	1	100	0	0	0	0
31	1	100	0	0	0	0	1	100	0	0	0	0
32	1	100	0	0	0	0	1	100	0	0	0	0
33	1	100	0	0	0	0	1	100	0	0	0	0
34	1	100	0	0	0	0	1	100	0	0	0	0
35	1	100	0	0	0	0	1	100	0	0	0	0
36	1	100	0	0	0	0	1	100	0	0	0	0
37	1	100	0	0	0	0	1	100	0	0	0	0
38	1	100	0	0	0	0	1	100	0	0	0	0
39	1	100	0	0	0	0	1	100	0	0	0	0
40	1	100	0	0	0	0	1	100	0	0	0	0
41	1	100	0	0	0	0	1	100	0	0	0	0
42	1	100	0	0	0	0	1	100	0	0	0	0
43	1	100	0	0	0	0	1	100	0	0	0	0
44	1	100	0	0	0	0	1	100	0	0	0	0
45	1	100	0	0	0	0	1	100	0	0	0	0
46	1	100	0	0	0	0	1	100	0	0	0	0
47	1	100	0	0	0	0	1	100	0	0	0	0
48	1	100	0	0	0	0	1	100	0	0	0	0
49	1	100	0	0	0	0	1	100	0	0	0	0
50	1	100	0	0	0	0	1	100	0	0	0	0













UNIVERSITY OF ILLINOIS-URBANA



3 0112 060296529